DOCUMENT RESUME

ED 195 274 IR 009 053

AUTHOR Jones, C. Lee
TITLE Linking Bibliographic Data Bases: A Discussion of the Battelle Technical Report.
INSTITUTION Council on Library Resources, Inc., Washington, D.C.
PUB DATE 15 Oct 80
NOTE 29p.: For related document, see IR 009 055.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Libraries: Cataloging: Cost Effectiveness: *Data Bases: *Information Networks: Interlibrary Loans: *Library Services: Methods: *Online Systems: Public Libraries: Reference Services: *Shared Services
IDENTIFIERS *Bibliographic Utilities: Linkage Analysis

ABSTRACT

This document establishes the context, summarizes the contents, and discusses the Battelle technical report, noting certain constraints of the study. Further steps for the linking of bibliographic databases for use by academic and public libraries are suggested. (FAA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED195274

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

LINKING BIBLIOGRAPHIC DATA BASES:

A Discussion of the Battelle Technical Report

by

C. Lee Jones

Program Officer

Council on Library Resources, Inc.
One Dupont Circle, NW, Suite 620
Washington, D.C. 20036

October 15, 1980

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

C. Lee Jones

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

LI200905 3

In 1979, the Council on Library Resources, Inc. (CLR) awarded a contract to Battelle-Columbus Laboratories to study the impact of linking the bibliographic data bases maintained by the U.S. bibliographic utilities. This document is intended to establish the context in which the study was carried out, to summarize its major points, to note certain constraints on the Battelle effort, and to suggest what the next steps might be.

Background

In the last ten to fifteen years there has been a significant improvement in the quality and quantity of bibliographic information available to library users. Printed and bound bibliographies of journal articles and books in specific fields or by individual authors are still published, but now often using computer technology as an aid to production. The text of the bibliography is prepared in machine-readable form, and the data base thus created is increasingly available as a supplement to the printed version. While in the past a printed bibliography was hopelessly out of date almost as soon as it was published, this no longer needs to be the case. If the work has been drawn from a machine-readable data base that is accessible and continues to grow, supplements to the printed version can be generated by the publisher or library user at will.

Success in providing information about information in this way has not been limited to one or two organizations. Technology has progressed to the point where any appropriately funded institution -- library, federal agency, private company, etc. -- can and has entered the field of bibliographic data base development. The National Library of Medicine was one of the first to begin to build data bases of citations to journal articles to support production of its printed products. The organizations that publish Chemical Abstracts and Biological Abstracts created machine-readable data bases for similar reasons. As these data bases grew, the advantages of searching them by machine became evident. Unfortunately, by this time each system had developed well beyond the point where it could be easily made compatible with the others.

The development of data bases of bibliographic records for books took a more fortuitous route. Aside from early independent efforts to produce machine-readable bibliographic records, the emergence at the Library of Congress (LC) of the MARC (machine-readable cataloging) communication format for bibliographic records dominated and made obsolete most other efforts to produce such records. As MARC records became available from LC, other institutions concentrated on developing computer processing systems to handle LC-MARC records. Initially, the goals were to take advantage of LC's cataloging work, to print catalog cards, and occasionally to support alerting services. But many computer system planners had more sophisticated ideas.

From these early efforts, the library user enjoyed only peripheral benefits, such as earlier receipt of books from library processing units and possibly the earlier appearance of catalog cards in the public

catalog. But the development of the Ohio College Library Center (now OCLC, Inc.) and BALLOTS (now the Research Libraries Information Network, or RLIN), along with efforts at the University of Chicago, the New York Public Library, the University of Toronto, the Washington State Library (Washington Library Network, or WLN), and elsewhere, promised much more to the patient library user. These systems, all built around the LC-MARC record, began to focus on services far beyond the production of catalog cards, such as shared cataloging, collection development, inventory control, circulation, interlibrary loan, serials control, and reference searching. Most of these services are still under development. Some, because of technological advances, are being implemented by local libraries independent of the institutions from which shared cataloging services are purchased. The result of all of this activity, particularly shared cataloging, is a set of rapidly growing data bases in many institutions, but especially at LC, OCLC, RLIN, UTLAS (University of Toronto Library Automated System), and WLN.

Consider the impact that these data bases and accompanying services have on the traditional methods of interlibrary loan (ILL). Using the OCLC service, which includes shared cataloging and an ILL subsystem, ILL personnel at an OCLC-served library can, within the limits of the data base, quickly examine the bibliographic records and codes indicating what other OCLC institutions own a desired item. A request for the item can be initiated through the computer terminal, and the lending library can respond quickly with information on the item's availability.

Though originally created to promote shared cataloging, it was not long before reference personnel were using the resulting data bases to answer a variety of reference questions. Some libraries have gone so far as to put terminals in the reference or card catalog area to promote the use of these systems for reference purposes. It is obvious to reference librarians that they now have available a very powerful reference tool. How much more powerful would it be if the several data bases were either linked or at least freely available for searching?

The advantages of bibliographic data bases and accompanying services to the library's technical services personnel are no less great. The library's catalogers can efficiently and quickly search for data supplied by LC and many other libraries and thus reduce the time needed to process a book before it can be used. These improved services lead to additional expectations by users and staff alike. If the single, computer-controlled data base can perform so well, why not develop the capacity to search and use the data bases (and even services) of other agencies?

In theory, the exchange of records among the shared cataloging systems or the linking of their bibliographic data bases would appear to be a simple matter. All of the data bases, for example, contain records based upon the MARC format, some produced by the Library of Congress and others produced by other institutions. Each shared cataloging service can accept records in the MARC format and most can also generate them for use outside the system. However, the processing systems that have been developed by each of these organizations are quite different; none are compatible. Each system was developed to exploit the most important

advantages offered by one or another data base management system, equipment vendor, or software supplier. Yet another impediment to the simple exchange of records is the complex nature of the MARC format. For two systems to recognize that they hold duplicate records, all data elements, including such things as commas and semicolons, must be identical. It is possible to produce intricate algorithms to assess the likelihood that two slightly dissimilar records represent the same item, but this is costly in terms of development and processing time.

The time may have come to bear that cost. In planning for future services, the library profession has developed the concept of a national bibliographic data base. Because of the separate and distinct operating systems involved in the development of bibliographic data bases, it is likely that instead of a single data base, the national data base will be a carefully articulated set of data bases. Philosophically, most of the library profession believes that this concept is a worthy goal. It has become evident that one way to achieve the goal is through a linking strategy.

In 1978 seven private foundations and the National Endowment for the Humanities provided funds to the Council on Library Resources, Inc. to establish the Bibliographic Service Development Program (BSDP), the objective of which is to help bring about a nationwide bibliographic record service in order to serve the needs and demands of the library user in the mid-eighties and beyond. Late in that year, the BSDP Program Committee began discussing the possibility of linking the U.S. organizations often referred to as "utilities," that is, OCLC, Inc., the Research Libraries Group's (RLG) bibliographic support system (RLIN),

the Washington Library Network, and the Library of Congress. The Program Committee, which includes the chief executive officers of the utilities, concluded that the discussion would be fruitless without some better sense of the economic and service impacts of such links. Among the questions raised were: Why should links be established? What types of links should be considered? Should different links be used for different purposes? What are the economic consequences of linking? No adequate data were available to answer these questions.

It was recognized that a thorough analysis of all possible links for the entire spectrum of potential services would cost too much in terms of time and money. The Program Committee agreed that a short-term study should be commissioned that would examine a fixed number of link alternatives for a limited number of services.

This was the situation when Battelle-Columbus Laboratories was selected to produce a technical report on the effects of linking the bibliographic utilities. Battelle began work on September 17, 1979, completing the study in May 1980. The final report, entitled "Linking the Bibliographic Utilities: Benefits and Costs", was submitted to CLR the following September.

Summary of Battelle Technical Report

The charge to Battelle was to evaluate the economic and service impacts of three link alternatives for three possible services: shared cataloging of current monographs, interlibrary loan, and reference

searching. The links were to be between LC, OCLC, RLG, and WLN. The approach used was to examine the effects of linking for each of the services and then to evaluate the costs of establishing three separate link configurations at each utility. The links examined were direct tape delivery, online "native" mode via dedicated line, and online "translation" mode via dedicated line. The precise nature of the link was ignored in the analysis of the impact of a link on the library services examined.

Link Benefits for Shared Cataloging

The benefits of linking for shared cataloging of current monographs were assessed by performing a hit-rate study. A group of items, selected from current cataloging activities at representative libraries plus a sample from the "Weekly Record" of publications contained in an issue of Publishers Weekly were searched against each utility's data base each week for eight weeks. A record was kept on each title reflecting when during the period various classes of records were discovered in the data base of each utility. The categories were: member-contributed records, CIP (Catalcging in Publication) records, and LC-MARC records. Once an LC-MARC record was found in a data base, the title was no longer searched on that utility.

The hit rates after eight weeks of searching were: 93 percent on OCLC, 87 percent on RLIN and WLN, and 86 percent on LC. If all four data bases had been linked, the combined hit rate would have been 96

percent. Clearly, in December 1979 and January 1980, LC, RLIN, and WLN had the most to gain from a link for current cataloging purposes.

The basis for calculating the cost benefit to libraries for linked shared cataloging was the assumption that for every record in a data base that could be found and used, the finding library could avoid original cataloging costs. Battelle constructed a "paper and pencil" model (composed of a sheet of data descriptors, calculation instructions, and blanks for data to be filled in by the model user) and, clearly stating the assumptions made in the model, calculated the economic impact of linking for shared cataloging. For all 2,100 OCLC libraries, annual costs avoided were calculated as $2,362,500, for all 80 libraries using RLIN $825,600, and for all 59 WLN libraries $227,445. As with subsequent paper and pencil models developed by Battelle, readers are encouraged to make their own assumptions, provide their own raw data and calculate the cost impact of linking for each service.

Link Benefits for Interlibrary Loan

First, the financial effect of linking for interlibrary loan purposes may be negative for some libraries, since their resources would be exposed to more libraries, which may result in more requests received and thus higher ILL costs. On the other hand, by taking advantage of such services, libraries may be able to verify more of the titles requested by their own users. Using a similar hit-rate study strategy for a sample of ILL transactions submitted by over 250 libraries, Battelle found that if all utilities were linked, an "average" OCLC

library could improve its verification hit rate by 3 percent. RLIN and WLN could have improved theirs by 35 percent.

Linked utilities would also improve the chances of finding a location for a desired item. In the same ILL study it was determined that all libraries could find a location for 79 percent of the items they requested if all utilities were linked. OCLC libraries would improve their location hit rate by 2 percent, while the RLIN and WLN rates would improve 46 percent through full linking.

Battelle again constructed a paper and pencil model to calculate the costs avoided through linking for verification and location purposes. Their assumptions are clearly identified so that different sets of assumptions can be tested by those interested in doing so. Using Battelle's assumptions the total annual costs avoided by all OCLC libraries is calculated as $22,880, for all RLIN libraries $38,350, and for all WLN libraries $13,216.

Link Benefits for Reference Searching

In order to assess the effect on the computing resource needs of utilities of linking for reference searching purposes, Battelle hypothesized that all local card catalogs would be replaced by an online union catalog maintained at each utility. This is a simplifying assumption for purposes of the study and is not a recommended course of action. The result is a worst case indication of computing resources required by each utility.

Though it was not possible to ascribe a value to the advantages of linked reference searching, Battelle developed another paper and pencil model to calculate the number of additional successful searches on one additional data base that might result from a linked environment. For this part of the study it was assumed that the card catalog was still in use and that only those searches, including subject searches, that "failed" at the card catalog were searched on the utility's union catalog. Then, only searches that failed at the utility's union catalog were considered eligible for linked searching. Given these limits, calculations indicate that for all libraries served by utilities, more than two million additional successful searches would be possible each year if the utilities were linked. While this analysis excluded searches intended to expand on or verify results from the local union catalog and those that might never be addressed to the local source in the first place, it does suggest that linking could result in substantial benefits to libraries and library users in the form of a large number of successful searches.

Link Alternatives Evaluated

In evaluating three link methods, Battelle used a network configuration of LC, OCLC, RLIN, and WLN. All four are assumed to use the same type of link. No combinations of links were examined, although the interactive computer model, BIBLINK, developed by Battelle for this project and used to handle large quantities of data, is capable of evaluating the cost consequences of a multi-type link configuration.

The links evaluated included direct tape delivery, which involves the exchange of records on tape among the utilities and represents the present state of linking. A "native" mode link was also examined. In this case, a utility makes it possible electronically for a member library to use its own terminal as though it were part of the other utility's network. The operator must then learn and use the unique commands and search strategies of the other utility. Much of the link burden is thus borne by the terminal operator.

The final link alternative is the "translation" mode, in which the terminal user enters search requests that are converted automatically by the library's utility into standard formats for transmission to the linked utility. Responses from the other utility are then converted by the library's own utility into its own formats. In this link option, the link burden (the translation) is accomplished by the computer systems on either end of the link.

Battelle's analysis of the three links resulted in the recommendation that online access using translation mode and dedicated leased lines would be the preferred link.

Implementation of a translation mode link is estimated to take three to four years and would require a substantial investment at each utility. The computer resources needed at each utility to support the added burdens of linked transactions are identified but will require additional analysis by the technical staff at each utility.

The computer processing and communications costs of supporting the translation mode link would be about $550,000 per year at RLIN and $270,000 per year at WLN. Because the computer processing costs at LC and OCLC are unclear, only the communications costs could be estimated -- $30,000 to $35,000 per year for each. Because inter-utility charges for service were not a subject for analysis, no assessment of incomes to utilities was made.

Recommendations

Following careful analysis of the data gathered, Battelle made these recommendations:

1. Continue to analyze link alternatives using the BIBLINK computer model.

2. Develop an online link with automatic translation of requests and responses.

3. Appoint two committees to establish standards, one in the area of communications protocols and one in the area of applications.

Discussion of Battelle Technical Report

Though some form of linking bibliographic data bases has long been assumed to be desirable, it is still useful to remind ourselves of the purpose of linking. There are several specific advantages to a system of links that, in essence, forges a de facto nationwide bibliographic data base out of the separate data bases that presently exist. But the single fundamental objective is to create an environment in which library users (scholars, students, business people, scientists, and all others) can examine a high proportion of the bibliographic records available in this nation, and ultimately beyond, and locate the material that is identified.

The Battelle study provides useful groundwork for moving toward this objective, for the investigators have gathered an impressive amount of data and have produced analytic tools that will be useful in the work ahead. As always there is considerable difficulty in assigning economic benefits to the enhanced services that would be available to library patrons, and Battelle was not asked to do this. However, links of various types are expected to have economic benefits for individual libraries. As discussed earlier, Battelle makes an effort to assess these benefits for the purposes of current monograph cataloging, certain interlibrary loan search operations, and reference searching. Clearly, links might produce other economic benefits for libraries in such areas as authority work, serials cataloging, retrospective conversion of

manual bibliographic records to machine-readable form, and even cooperative acquisition programs. These activities were not included in the Battelle study because, in addition to factors of high cost and lack of time, CLR and the BSDP Program Committee felt that the three services that were studied would provide a fair indication of the costs and service impacts of linking.

The direct costs of developing, establishing, and maintaining a system of links would affect individual libraries, since the increased costs for linked services would be passed on to them by the utilities being used. Some believe that the economic impact on the utilities is the more important issue. But utilities are service agents, reservoirs of bibliographic data and service, maintained for the benefit of library users. Links should not be evaluated solely on how they affect the utilities but also on how they affect the libraries served by the utilities. Costs and potential service degradation (e.g., increased response time) are effects that are important to both libraries and utilities, but library managers, not utility managers, ought to have the responsibility for making the final judgment on linking. It is also clear that utility managers have an obligation to explain to libraries the anticipated changes in cost and service that would result from proposed links.

The problem with addressing these suggestions to the utilities is similar to one of the problems faced by Battelle during the study. One of the more common misconceptions is that the organizations popularly referred to as "utilities" are identical organizations in different stages of development. In the Library of Congress document Network

Planning Paper #2: A Glossary for Networking, the phrase "bibliographic utility" is defined as "an organization that maintains online bibliographic data bases, enabling it to offer computer-based support to any interested user...." Of the four utilities that Battelle studied, all would qualify as utilities according to this definition with the exception of the Library of Congress. LC does not maintain an online data base for remote access, but it does distribute its data base to its users, which include the other utilities. The bibliographic systems operated by OCLC, RLG, and WLN fit the definition, but they are very different organizations.

OCLC, Inc., provides services from its data base principally through state or regional network brokers. These organizations handle marketing, training, and other services developed by the broker. OCLC also serves some libraries directly. While other utilities view themselves as cooperative operations directly responsible to members or owners, OCLC, at least among its employees, views itself as a company with corporate goals that cannot match all of those of the diverse institutions it serves. OCLC invites all types of libraries to take advantage of its services; it has users in every state of the Union, in Canada, and in Mexico and will soon begin offering services in western Europe.

RLG, while committed to providing bibliographic services to those institutions that used the old BALLOTS system, is a "program-driven" organization, which uses its bibliographic utility, RLIN, to serve the needs expressed by its owner-members. The utility portion of the RLG operation is simply a tool to help achieve other program goals. RLG

does not intend to serve the needs of every kind or size of library. It designs its programs and activities to meet the needs of large research libraries, which are its membership targets. Thus, while there are no geographic limits to RLG, there is a "type-of-library" limit.

WLN is more akin to a cooperative, with members helping to shape the development of additional services. WLN also has program activities not necessarily related to its bibliographic utility. The Pacific Northwest is the area defined for WLN operation. In that sense alone, it is similar in size and service demands to other regional networks that broker OCLC services. WLN welcomes as members any type of library within its geographic boundaries (Alaska, Washington, Oregon, Idaho, Montana and the western provinces of Canada).

The only obviously common element among these three utilities is the presence of online data bases that provide the foundation for a variety of services. In spite of this basic similarity, Battelle was faced with a substantial anomaly. OCLC and WLN, though growing steadily, had each achieved relative stability in the services offered and in market served. RLG, on the other hand, was right at the base of what appeared to be, and subsequently became, a spurt of rapid growth. The data collected by Battelle on RLG from November 1979 through January 1980 represented only three full and two affiliate members, and the users that had been part of the pre-RLG BALLOTS system. Data collected on RLG for use in the BIBLINK model (July 1980) reflects thirteen full members. Now, as the Battelle report is issued, RLG has twenty-one full, three associate, two affiliate, and several special members, and, of course, the original BALLOTS users. Battelle's response to such a

dynamic situation was to develop analytical tools that can be used over and over again, as conditions and assumptions continue to change. Consequently, although the data displayed in the report is valid only for the brief time period studied -- November 1979 through February 1980 -- we now have the tools to assess conditions as they exist now or in the future.

In defining for Battelle the library services that might benefit from linking, CLR and the BSDP Program Committee selected those services that were offered by all the utilities. Shared cataloging has from the beginning been the backbone of the bibliographic utilities, providing needed revenue and enhancing the size of the data base.

Shared Cataloging

In order to assess the benefit of linking the utilities for shared cataloging, Battelle performed a hit-rate study for current monographs. The study contained valid data on English-language monographs. Since only 12 percent of the sample were non-English titles, broad conclusions on this category may be tenuous, though indicative. Non-current monographs were not a part of the Battelle study, so no information was generated concerning the effect of linked utility data bases on retrospective conversion projects.

Because of subjective decisions that vary from institution to institution, the issue of the quality of records available for sharing was not addressed in the Battelle study. It is well known that libraries will not accept records from certain other libraries because of a perceived "lack of credibility" of their cataloging expertise.

Most libraries accept LC copy as a known quantity with a recognized and consistent level of quality. A subsequent cataloging hit-rate study might profitably identify LC-source records (those drawn by other libraries from National Union Catalog records, for instance) so that the original source of a record might be identified.

Some libraries, wishing to avoid the high costs of original cataloging, choose to hold materials for months in hopes that the needed record, preferably an LC record, will become available. The Battelle study seems to indicate, however, that for current imprint, English-language materials, it does not pay to wait more than four to six weeks for copy. But this concept of "waiting for Godot" flies in the face of one of the basic tenets of utilities operated for shared cataloging purposes. In order to make a shared cataloging service work to maximum effect, libraries must generate an original bibliographic record if one is not available to share. This tenet is difficult to enforce but one that helps make shared cataloging an effective cost reduction strategy.

Interlibrary Loan

In the area of public services, Battelle assessed certain effects of linked searching for verification and location information for interlibrary loan purposes. There are many other elements that were not a part of this study but nevertheless impinge upon the quality of interlibrary loan services. It is clear that ILL fill rates, though affected by the availability of location information, are far more severely influenced by institutional policies on filling loans and inter-institutional agreements that provide for preferential treatment

for cooperating partners. There are also costs to individual libraries that will grow out of a system of links between utilities for ILL purposes. Each institution's holding records would be available to a much wider audience resulting, quite possibly, in higher levels of ILL requests received. The costs of handling these additional requests are related to the link. Library managers must then measure the benefits they derive from the link and decide how to handle the increased costs -- within the budget as a cost of access to information about other collections or by charging access fees to those requesting service by virtue of the link.

So far, among the three utilities, only OCLC has an online ILL request service, although RLG plans to implement one. Links between the utilities would create a situation in which OCLC might provide ILL request service to non-OCLC institutions at a price premium that could help offset part of the cost of the link and help pay for work created at the library from which material was being requested. Some institutions might find the payment of a premium under these circumstances a relative bargain given other less efficient options.

Reference Searching

The study suggests that there are substantial advantages to the individual library patron to be derived from linked utility data bases. The direct benefits are in the form of expanded reservoirs of information for interlibrary loan and reference searching purposes. The Battelle study makes no effort to quantify these benefits, but much of the basic data are available for such an analysis.

The analysis of the quantity of reference searching that might be done if the data bases were linked is based upon relationships between circulation and use of the card catalog established in other studies. Some of these have been validated and so lend weight to the Battelle conclusions. The advantages to library users served by libraries in any network are likely to be substantial. They are likely to be especially valuable to OCLC library users in the area of subject searching, a service not yet offered by OCLC. But any library user seeking a comprehensive search, whether by subject or by some other access point, will find that access to the principal bibliographic data bases most helpful.

Selection of Options

The strength of the Battelle effort is in the data collected and in the development of tools to display and analyze the data. The evaluation of specific types of links, three carefully selected of the more than four hundred possible configurations, is the starting point for the assessment of other alternatives. If each utility and each service offered by the utility required a different link, the number of potential link combinations would run into the many thousands. However, many options can be immediately eliminated for one or more practical reasons. Costs may be too high, for example or the resulting service too duplicative, unresponsive, or so cumbersome as to discourage use. Chapter Four of the report includes an extended discussion of link options.

In the selection of links to be examined, it was decided that only direct physical links between existing organizations would be evaluated. It was recognized that a central clearinghouse might be a viable alternate link mechanism, but this alternative was beyond the scope of the present study. It was also recognized that passive links, that is, links that did not require major software development but merely policy changes, were possible. In such a scenario, each library using any one of several terminals would select the appropriate utility for the activity to be performed. A library might first search its "host" utility, expanding the search to other utilities by using TELENET or TYMNET to dial the one or more additional utility data bases to be searched. As with any link alternative there are a great many problems to be resolved in such a passive link environment.

Options Explored

The three link options actually explored by Battelle were selected for a variety of reasons. The direct tape delivery was selected because it approximates the nature of the limited linking that now takes place. The MARC family of data bases are distributed by LC in tape format. So far, no institution or utility is online to LC in order to receive newly generated MARC records. However, the MARC-S (Serials) records are created online by LC on the OCLC system. Similarly, the CONSER data base, maintained by OCLC, has been duplicated in Columbus and sent to LC for distribution on tape to subscribing utilities and libraries. This link form is awkward and expensive but it does exist and represents a form of link in the broadest sense.

The native mode and the translation mode, each via dedicated lines, were selected for analysis because they apparently would be efficient, direct physical links of the utility data bases. Both are complicated and required the sophisticated manipulation of many cost elements by the computer model, BIBLINK, developed by Battelle in order to assess the economic impact of each link mode. The model, given appropriate cost data, can be used to assess a number of other linking options, and with this information the linking discussions can proceed.

The Battelle recommendations are based upon the link alternatives that they studied. Recognizing the limitations of their analysis, they urge that the tools developed in the study, especially BIBLINK, continue to be used in efforts to analyze other link alternatives and to reassess the alternatives examined in the present study as new data becomes available. (Verification of the accuracy of cost data from the utilities would be particularly valuable in confirmation studies.) Of the link alternatives studied, the translation mode, because of the ease of use by the terminal operator, was believed to have advantages over the native mode. Battelle recommends that work be started to develop the translation mode link.

Battelle also recommends the formation of two committees: an applications standards committee and a communications standards committee. Elements of these committees are in place and operating. The communications standards committee, in concert with the appropriate American National Standards Institute subcommittee, might be the WLN/RLG/LC committee currently working on the telecommunications plan and standards for the CLR-supported Linked Authority Systems Project.

The applications standards committee could be considered to be already at work in the form of the BSDP Task Force on a Name Authority File Service. True, the latter committee is limited to one application (authorities) at the moment, but careful additions to this group might well bring into existence the sort of committee Battelle believes is necessary.

The Future for Linking

The conclusion of the Battelle linking study signals the beginning of a period in which the utilities and the libraries they serve will design the first of what will likely be a series of links between data bases. In areas where technology (both computing and telecommunications) is changing so rapidly, it is unlikely that the first link choice will prove to be the final one.

Battelle has examined three link options. Must we examine all other options? Unlikely. But those that appear to have promise should be examined, including any that appear to achieve the objective at reasonable costs. Costs that may appear unreasonable from one perspective, for instance the operating utility, may have such dominant user service benefits that ways must be found to bear the costs. The benefits must ultimately accrue to the individual library user in terms of expanded service or additional information. Ways must be found to support the added costs of these benefits possibly by redeploying

present fiscal resources. If we accept the premise that the benefits of linking are substantial, and the partial evidence generated by Battelle suggests that they are, then we must find ways to realize and finance these benefits. However we accomplish this, it will only come into existence through cooperative efforts on the part of utilities, participating libraries, and funding sources.

It is important to keep in mind that the entire operating environment is changing. Data bases will continue to grow at enormous rates, new services will continue to be explored and developed, and new ways of delivering information to new sites will continue to be created. Thus, some basic, though speculative, questions will eventually need to be asked. Is it reasonable, for instance, for utilities to maintain the entire LC-MARC data base online? To maintain records that are seldom or never used? To maintain records that have been used only once or twice and not at all in the last year? To maintain specialized data bases that are used only infrequently? We may be approaching the point at which most of these questions will be answered by the utilities in the negative. After all, there are costs associated with maintaining little-used records, which must be borne by all participating libraries in the present service-pricing structure.

The concept of a central pool of little-used bibliographic records has been discussed before and discarded as an expensive alternative to storage costs. Storage costs continue to decline, further reducing the appeal of a central pool. But it is possible that there are substantial advantages to creating a central pool containing every record maintained

by the utilities. This would create a true national bibliographic data base and would provide opportunities that do not now exist.

It has been said many times that utilities try to provide two very different services that are incompatible: the development of a data base and the retrieval of information from that data base. In order to develop the data base efficiently, certain compromises must be made that degrade the potential efficiency of the retrieval functions. In attempting to provide the best of both worlds, the utilities have worked particularly hard to strike a balance between these conflicting functjons. Compromises have been made and will continue to be made.

A central pool of all bibliographic records that also provided a sophisticated information retrieval service would allow the utilities to concentrate on data base development and other support services. The major deterrent to this scenario is the enormous resources already committed by utilities to the development of retrieval systems. If the pool were used by all utilities, linking might only be required between the utility and the central pool. If it still were necessary to link among utilities, the link might take place through the central pool. There are those who have discarded the notion of a central pool or clearinghouse, but it seems reasonable to explore this option in greater depth with attention paid to how such a facility would be managed and by whom.

Finally, we ought to explore the possibility of leaving the choice of utility for a particular service up to the individual library. Clearly, for shared cataloging, only OCLC can be considered at present as everyone's utility, since both RLIN and WLN have limitations on

institutions eligible for their shared cataloging service. However, for reference searching and for interlibrary loan functions, a system that allowed libraries to select the desired data base might preclude the need for links between the utilities themselves. Even authority work would be much simplified when the Name Authority File Service becomes available if institutions had the capacity to access the file directly from any utility-supported terminal.

Next Steps

No matter which option is finally selected to provide library users with access to the bibliographic work of the nation's libraries, it will require some sacrifice of autonomy on the part of individual libraries and utilities. We must be bold enough to accept such sacrifices in order to provide the services that an increasingly sophisticated library public will demand.

Based on this discussion, the next steps for network planners and managers appear to be:

1. Convene a meeting of members of boards of governors (directors, trustees, etc.) of utilities to discuss linking and alternatives. Attempt to identify links desirable to all utilities, their libraries, and library users.

2. Continue to work toward the development of standards that will make possible a variety of link alternatives.

3. Use BIBLINK to evaluate other promising link alternatives.

4. Investigate alternatives to linking the utilities directly to each other, such as direct terminal access to any desired utility or creation of a central pool of bibliographic records.